[17] P. Parent and S. W. Zucker, *Trace Inference, Curvature Consistency, and Curve Detection*, IEEE Trans. PAMI, Vol. 11, No. 8, Aug. 1989, pp. 823-839.

[18] B. Parvin and G. Medioni, *A Dynamic System for Object Description and Correspondence*, Proc. CVPR, Jun. 1991, Maui, Hawaii, pp. 393-399.

[19] P. Perona, *Steerable-Scalable Kernels for Edge Detection and Junction Analysis*, Proc. ECCV, May 1992, Santa Margherita Ligure, Italy, pp. 3-18.

[20] I. Rock and S. Palmer, T*he Legacy of Gestalt Psychology*, Scientific American, Dec. 1990, 84-90.

[21] H. Rom and G. Medioni, *Hierarchical Decomposition and Axial Shape Description*, IEEE Trans. on PAMI, Vol. 15, 1993, pp. 973-981.

[22] A. Sha'ashua and S. Ullman, *Structural saliency: the detection of globally salient structures using a locally connected network*, Proc. ICCV, Dec. 1988, Tampa, Fl., pp. 321-327.

[23] F. Stein and G. Medioni, *Recognizing 3D Objects from 2D Groupings*, IUW92, Jan. 1992, San Diego, Ca., pp. 667-674.

[24] K. Sugihara, *An Algebraic Approach to Shape-From-Image Problems*, Artificial Intelligence, 1984, Vol. 23, No. 1, pp. 59-95.

[25] S. Ullman, *Filling-in the Gaps: The Shape of subjective Contours and a Model for Their Generation*, Biological Cybernetics 25, pp. 1-6, 1976.

[26] F. Ulupinar, *Perception of 3-D Surfaces from 2-D Contours*, IEEE Trans. on PAMI, Vol. 15, 1993, pp. 592-597.

[27] D. I. Waltz, *Generating Semantic Descriptions from Drawings of Scenes with Shadows*, Chapter 3 in The Psychology of Computer Vision, P. H. Winston (Ed.), Mc-Graw Hill, New York.

[28] M. Wertheimer, *Untersuchungen zur Lehre von der Gestalt, II. Psychologische Forrschung*, 1929, 4, pp. 301-350. (English translation)

[29] L. R. Williams and D. W. Jacobs, *Stochastic Completion Fields: A neural Model of Illusory Contour Shape and Salience*, Proc. of the 5th Int'l Conf. on Computer Vision, Cambridge, Mass. 1995.

[30] S. W. Zucker, A. Dobbins, C. David, and L. Iverson, *The Organization of Curve Detection: coarse Tangent Fields and Fine spline Coverings*, Proc. ICCV, Dec. 1988, Tampa, Fl., pp. 568-577.

curves ending abruptly. Since all computations are performed on a discrete grid, quantization and rounding errors restrict the selectivity and amount of clutter the system can handle.

It is possible to base much of the reasoning of the voting and interpretation on a firmer statistical footing. In a recent work, Williams and Jacobs [29] have done just that. We opted for a more intuitive way to derive the theory.

References

[1] N. Ahuja and M. Tuceryan, *Extraction of early perceptual structure in dot patterns: integrating region, boundary, and component Gestalt*, CVGIP 48, 1989, pp. 304-356.

[2] D. H. Ballard, *Generalizing the Hough Transform to Detect Arbitrary Shapes*, Pattern Recognition 13, 2, 1981, pp. 111-122.

[3] K. R. Boff, L. Kaufman, and J. P. Thomas, Handbook of Perception and Human performance, Vol. II, John Wiley and Sons, pp. 36-1 - 36-30.

[4] M. B. Clowes, *On Seeing Things,* Artificial Intelligence 2, 1, 1971, pp. 76-116.

[5] J. Dolan and R. Weiss, P*erceptual Grouping of Curved Lines*, Proc. IUW89, Palo Alto, CA., pp. 1135-1145.

[6] R. O.Duda and P. E. Hart, *Use of the Hough Transformation to Detect Lines and Curves in Pictures*, Communications of the ACM, Vol 15, 1972, pp. 11-15.

[7] G. Guy and G. Medioni, *Perceptual Grouping using Global Saliency enhancing operators*, Proc. of ICPR92, The Hague, Holland, 1992, pp. 99-104

[8] G. Guy and G. Medioni, *Perceptual Grouping using Global Saliency enhancing operators*, IRIS-USC Technical report.

[9] F. Heitger and R. von der Heydt, *A Computational Model of Neural Contour Processing: Figure-Ground Segregation and Illusory Contours*, 1993 Proc. of the ICCV, pp. 32-40.

[10] H. Helson, *The Fundamental Propositions of Gestalt Psychology*, Psychological Review, 1933, 40, pp. 13-32.

[11] P.V.C. Hough, *A Method and Means for Recognizing Complex Patterns*, U.S. Patent No. 3,069,654, 1962.

[12] G. K. Kanizsa, *Subjective contours*, Scientific American, April 1976.

[13] J. J. Koenderink, Solid Shape, The MIT Press, Cambridge, 1990.

[14] D.G. Lowe, *Three-dimensional object recognition from single two-dimensional images*, Artificial Intelligence 31, 1987, 355-395.

[15] R. Mohan and R. Nevatia, *Segmentation and description based on perceptual organization*, Proc. CVPR, Jun. 1989, San Diego, Ca., pp. 333-341.

[16] R. Mohan and R. Nevatia, *Using Perceptual Organization to Extract 3-D Structures*, IEEE Trans. on PAMI, Vol. 11, No. 11, November 1989, pp. 1121-1139.

eccentricity measure at each site. Table 2 summarizes the main differences and similarities of the two methods.

Table 2: Comparison between the Hough Transform and Our Scheme

	Our scheme	Hough Transform
Complexity	Same for all shapes and properties	Grows with the number of parameters
Space Requirements	Constant (*O(image size*))	Grows with the number of parameters and resolution.
Properties (Families of shapes) coding	Yes, by defining suitable fields	Not in any obvious way
Analytical shapes	Yes, by the same method	Yes
Voting pattern[a]	2-D	1-D[b]
Localization of shape	Yes	Not always. Further processing is needed to find location
Means of output	A ridge of maxima	A peak in parameter space
Choice of bin size	Always image resolution	Hard (May be crucial)
Selectivity	Low (Because of compaction)	High
Transformation	Non-Linear	Linear

a. Our nomenclature. Refer to text for definition.
b. Or at least one dimension less that of the problem.

7 Summary and Conclusion

We have introduced a unified way to extract perceptual features in edge images. By ‘unified’ we mean that all low-level features (edgels, points) are treated in a uniform way, and no special cases exist. The scheme is threshold-free and non-iterative. It is especially suitable for parallel implementation, since computations of the saliency maps are independent for each site, and parallel algorithms for line following are known and can easily be adapted. Also, calculations are simple and stable, as no curvatures or any other derivatives need to be computed on the digital curves.

The system can rank features based on their perceptual importance. This allows a real-time application to process as many features as time permits.

Some of the issues which have not been addressed are the resolution dependency of the description. At this time, only one level of description is possible. Also, we have not tried to localize end-points of

For more complex shapes, the scheme can be generalized by increasing the dimensionality of the accumulator space. For instance, circle detection requires 3 dimensions (2 for the center, 1 for the radius).

The next phase in the process is to search for peaks in the parameter space array. The desired feature is only defined in terms of the above parameters and further processing is needed to localize the actual segment in the image plane. Note that this process is absolutely global, as it ignores distances between contributing candidates. If some orientation data is known at a site, fewer candidates gather votes, and the ability to reveal the desired shape gets better.

The same scheme can be extended for other shapes [2] by extending the parameter space.

6.2 Our scheme as a Hough transform

Our system can be viewed as a Hough Transform where *the parameter space is the image plane itself*. This allows for many more degrees of freedom in the choice of shapes, and in the basic definition of the desired shapes. This choice of the parameter space allows us to define other **voting patterns**[1] which enable us to encode the constraints (see 2.1.1).

Our scheme is thus capable of finding shapes described by their *properties* (smoothness etc.) rather than by their *exact analytical parameters* (as in the Hough Transform).

Note that the classical Hough transform for edgels can be implemented in our scheme. We replace the Extension Field by a straight line, and remove the attenuation factors. For non-oriented inputs, the Point field (again without the attenuation) will behave in the desired fashion. Also, note that if our voting method is kept, we can detect junctions at the same time.

To implement a circle finder using our scheme, we would use our Extension Field, but with all field strengths set to the value 1. Coding of the *other* constraints[2] is where the strengths of the individual field elements come into play. *This is not possible with the original Hough Transform*.

The down side is that the result is not found at an isolated peak of the parameter space, but rather as a continuous ridge of peaks (in the case of the Extension Field). Also, note that the parameter field we use is *always* the image plane, which, in many cases, is smaller than a classical Hough parameter space. This ‘Compaction’ of data is the cause for the low selectivity compared to Hough transforms.

However, for Perceptual Grouping, this loss of selectivity seems to be an advantage. Note that the Hough Transform could find clusters(e.g. straight lines) even when they are *not* salient, because the notion of interfering features is not made explicit. This makes the Hough Transform a *linear superposition.* Our scheme, on the other hand, is highly non-linear and takes into account the interference by computing the

1. Similar to the sinusoid defined for the straight line detector
2. that make it a perceptually appealing scheme

in Figure 23. This kind of field will vote strongly for straight angle T-junctions, and weakly for other di-

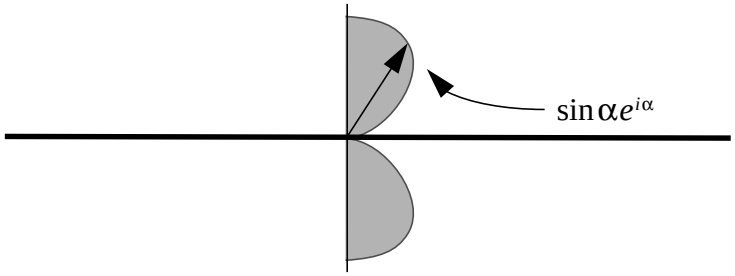
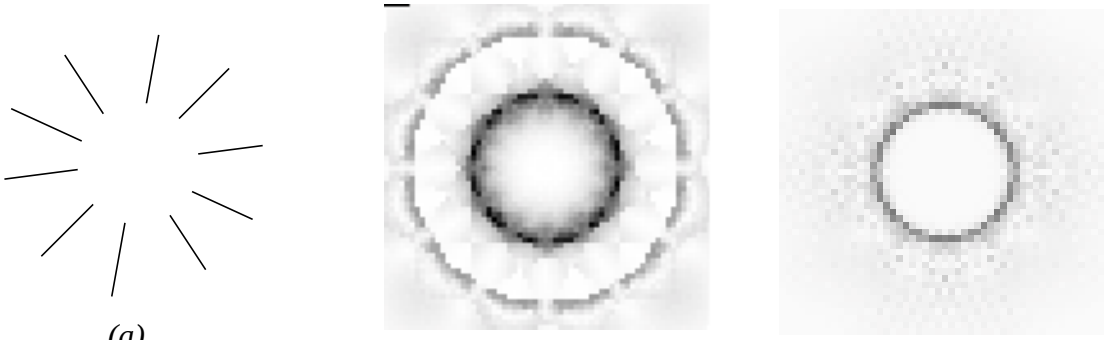


Figure 23 A multi-directional edge constructor for the End-Point Field. (envelope shown)

rections. This merely means that more support is needed for more acute angles.

5.4 Experimenting with the End-Point field

We tested the end-point field with the synthetic image in Figure 20. It is clear that the outer circle



(a) *(b)* *(c)*

Figure 24 Results of applying the end-point field. Note that the outer circle is not highlighted! (a) Original image. (b) Saliency map. (c) after thresholding single votes.

receives relatively low saliency, and can be completely removed if we consider single votes in the plane as wrong hypotheses (see Figure 24).

6 Comparison With the Hough Transform

It is interesting to contrast our scheme with the familiar Hough transform. In the next sub-sections we show, that in a sense,

6.1 The classical Hough transform

Consider he classical Hough transform [6,11] which can detect co-linear clusters in a *dot image* in the presence of noise. This is accomplished by detecting peaks in an accumulator array. If the line is parameterized in terms of (θ,d), each point votes in θ,d space for a sinusoid with equal weights. The votes in each cell are scalar, and the accumulation is simply a sum. This scheme is trivially extended to oriented edges, in which case, each edgel votes for a cell.

The same result holds for the 3-D case, since the distribution of the projection (either orthographic or perspective) of any line to 2-D is equal to the marginal distribution of the original 3-D distribution. That is, uniformly distributed segments in space project to uniformly distributed segments in the plane.

Piecewise smooth curved objects can be approximated by linear segments to any degree of accuracy, and should thus satisfy the same distribution. The same is true for lines of different lengths. It is easy to see that the length of the segment does not change the probability of a given intersection angle.

5.2 Convex T-junctions are more common

We believe that the a priori probability of having convex T-junctions is larger than concave T-junctions. This observation can be illustrated through a simple example, as shown in Figure 22. Figure 22(a)

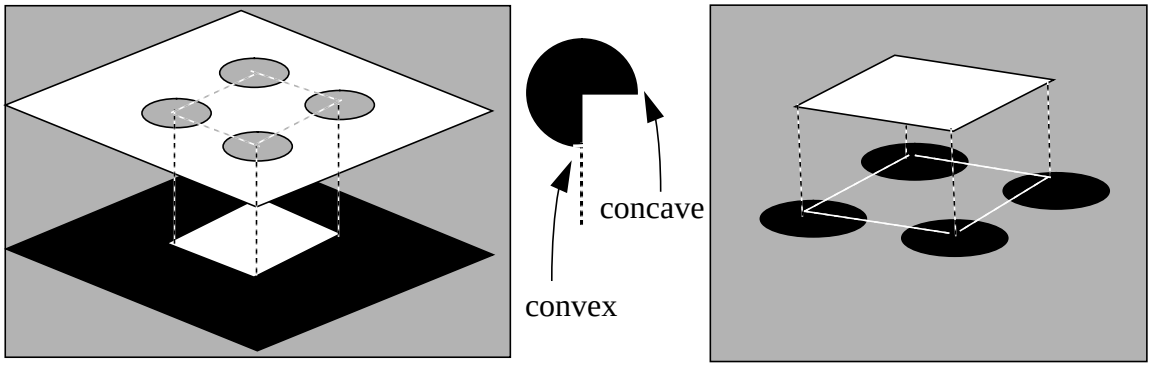


Figure 22 Convex and concave T-junctions. The Kanizsa square has two valid 3-D interpretations, but only one of them is perceived (b). The other one in (a) requires us to imagine concave T-junctions.

depicts a physically valid interpretation of the Kanizsa square, but one that is rarely perceived. We believe the reason is that such an arrangement of circular disks and a square gives rise to concave T-junctions. Figure 22(b) however, gives the more commonly perceived arrangement.

Figure 24(a) is another example where the inner circle is perceived, but the outer one is not. Again, we claim this is because concave T-junctions are not normally hypothesized.

Many other experiments (such as in [3]) indicate that end-point formations that require concave T-junctions are not normally perceived. Also, real objects tend to have longer convex boundaries, than concave ones (see [13]), and since boundary length is proportional to the probability of an intersection, convex T-junctions are more common. The above observation provides an additional constraint when designing an End-point field, as shown next.

5.3 Building the End-Point Field

The angle distribution derived previously suggests convolving our original Extension field with a multi-directional edge having a diameter function of a sinusoid, as was done for the point field, and shown

5.1 Straight angles in T-Junctions are more likely than any other

We derive the distribution of T-junction angles in a random world and show that close-to-straight T-junction angles are more likely to appear than any other angle. We claim that our human perception uses this property, and thus attempts to perform perceptual tasks on end-points stimuli only when the illusory intersection angles justify it. In Figure 20(a) all lines meet the illusory contour with almost straight angles, thus making the shape 'visible'. In Figure 20(b) the lines are occluded by an exact circle, but the angles are much more acute. *No* perception of shape is evident, even though the end-points trace an exact circle.

Claim: Given two unit size segments, we independently drop each one of them on a finite board, with uniform probability with respect to position and angle. We then look at the distribution of the intersection angles between the two segments. We claim that angles close to 90 degrees are more likely to appear then acute intersection angles. (We do not count cases where the segments do not intersect.)

Proof: Without loss of generality we assume that the first segment lies at the origin of the plane. Since all positions in the plane have equal probability, the probability of intersection is proportional to the area of the locus of positions *allowing* intersection for a certain angle. This is obviously a parallelepiped, as shown in Figure 21. The area of this parallelepiped is proportional to **sin**α for α between 0 and 90 de-

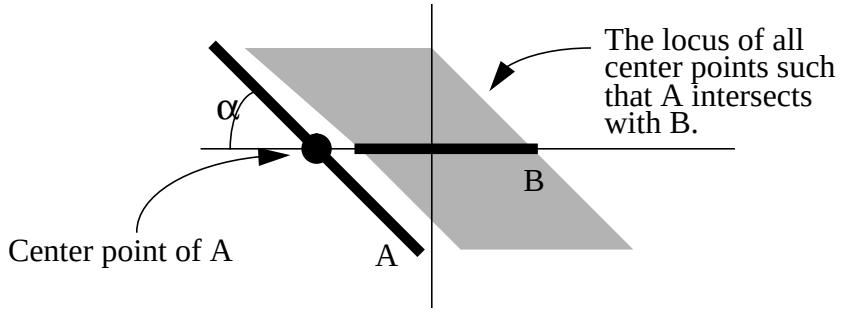


Figure 21 Given two segments A and B, the grey area is proportional to the probability that A and B intersect for a certain angle α.

grees. We can thus write the distribution function of intersection angles as:

$$F(\alpha) = P(x \le \alpha) = \begin{cases} 0, \alpha < 0 \\ \int_0^{\alpha} \sin x dx = 1 - \cos\alpha, 0 \le \alpha \le 90 \\ 1, \alpha > 90 \end{cases}$$

Clearly, angles close to 90 degrees are more probable. ❑

Adding more lines to the system will not change the distribution of angles, since we can consider all segments to be pair-wise independent.

2% noise 4% noise 6% noise Original with 6% noise

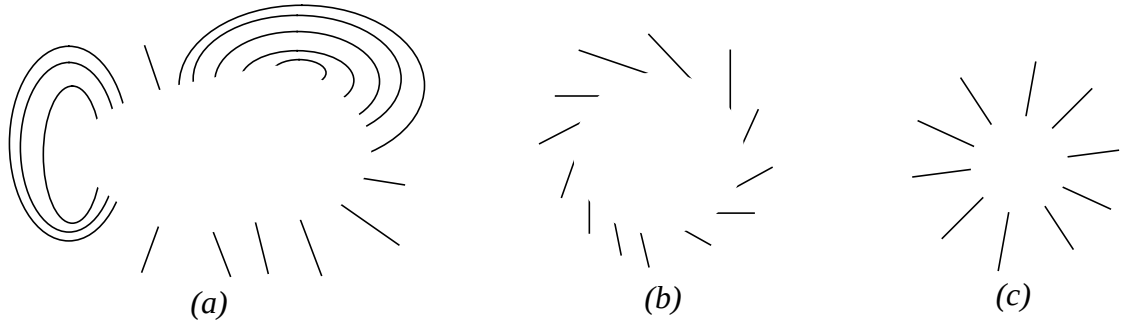
Figure 19 Local maxima of the saliency maps, for different noise levels.

4.6 Complexity Issues

A naive way to implement the algorithm requires $O(n^2k)$ operations, where *n* is the side size of the image, and *k* is the number of edge elements in the input image. In practice, the local density of edgels restricts the useful scope of the field. Alternatively, instead of computing a *dense* saliency map, we can compute the saliency of existing edgels only. Here we assume the existing edgels are stored in some condensed array, so that traversal of the whole image is not necessary. This results in complexity of $O(k^2)$, and can be useful as a focus of attention map. This mode allows then for a second pass on the salient features only.

5 Application of scheme to end points (End-point field)

In many cases, we tend to interpret end-points as being a partially occluded line. If enough end-points are available, they support the hypothesis of a shape occluding a collection of lines. This is a well documented phenomenon, and has been partially implemented by Heitger and von der Heydt [9], in the case where end-points are already marked. Illusory edges appear to outline the said occluding shape. Figure 20 illustrates end-point scenarios.

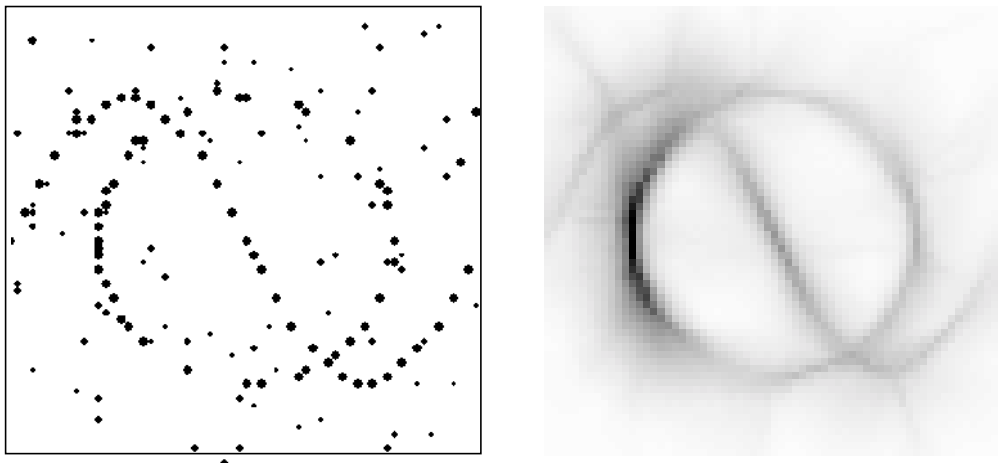


(a) *(b)* *(c)*

Figure 20 An end-point formation. (a) A center egg-like shape is not only perceived but also looks whiter (after [3]). (b) An invisible circle occludes lines. No sensation of a circle is evident. (c) The inner circle is perceived, but the outer one is not!

4.5.2 Point Field

We tested our system on the image in figure 18(a). Initially, the system was run using the Point field. This resulted in a saliency map *with* orientation data. A second phase of computation was then performed, using the directional Extension field (Figure 6). That stage produced the final saliency map as shown in figure 18(b).



(a) *(b)*

Figure 18 (a) A non-directional input image. (b) Saliency map, after applying the Point field and the directional extension field

4.5.3 Noise Breakdown Point

Our algorithm degrades gracefully and fails in noise, unlike the Hough transform, for example, which may find line formations even in a very cluttered image, where such formations are not visible, and could be a result of accidental alignments.

We have performed a controlled experiment to determine the amount of random noise which still allows for a correct following of a curve, given a constant factor of existing edge. The test is whether all points along the perceptual circle belong to the local maxima, as defined before. Only uniform noise in space and orientation was applied to the original image. The circle has a radius of 25 pixels and consists of 33 segments of unit strength and the percentages of noise applied are 2,4, and 6 percent[1], again of unit strength. This means that about a quarter of the circle edge exists. The local maxima of the corresponding saliency maps are shown in Figure 19. Up to 4 percent of noise, the circle (or large parts of it) is recoverable. With 5 percent noise and more, the saliency map degrades, and it is no longer possible to apply the following algorithm. The original image with 6% of noise is also given in Figure 19.

1. that is, each pixel has a 0.06 probability of becoming an erroneous site.

tropic and cannot affect the direction of true votes in any area. It can be seen that close-by areas get low agreement values (Figure 15(a)), and far away areas get low voting weight(Figure 15(b)). In both cases, the interference is small. This is true for all smooth curves in the image.

Noise Contamination Random segments *do* contaminate the environment, and their effect is reduced through the robustness of the voting scheme. In the experimental results section, we empirically test for allowable levels of noise.

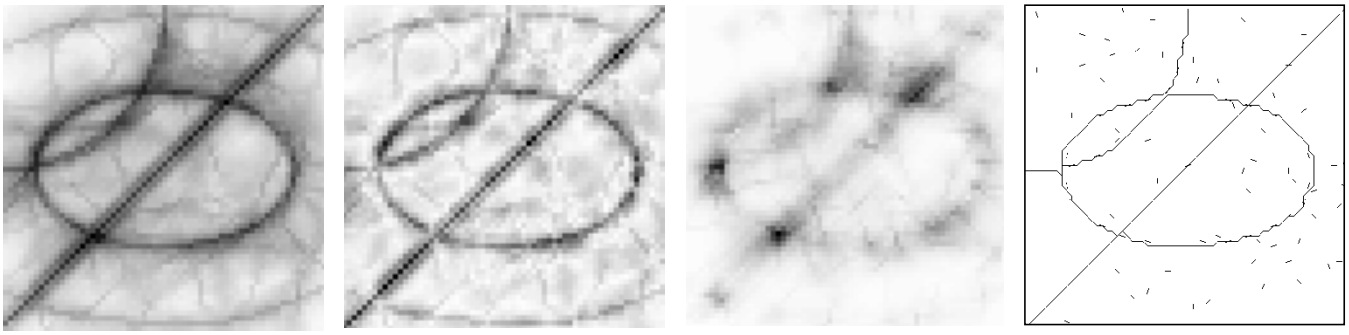
4.5 Results

4.5.1 On Synthetic Images

We have tested our approach with the synthetic data shown earlier in Figure 1.The saliency map produced is shown (strength only) as a grey-level image in figure 16 and the result of following the path of highest saliency produces a “clean” circle. Figure 16 also shows the result of the same procedure for the other scenes. Figure 17 shows an example of the steps involved in producing a high-level description of a given image, using the junction map in conjunction with the saliency map.

(a) *(b)* *(c)*

Figure 16 The Raw Saliency maps of images in figure 1.



(a) *(b)* *(c)* *(d)*

Figure 17 Extracting the most salient features. (a) Largest eigenvalue strength map. (b) Eccentricity enhanced map. (c) Junction saliency map, and (d) linking.

4.4.2 The Non-maximum Suppression Phenomena

We have mentioned the superior selectivity of the Enhanced Saliency map. To illustrate this behavior, we look at the eccentricity *only* map of a straight line (Figure 15(a)). Note how low the eccentricity

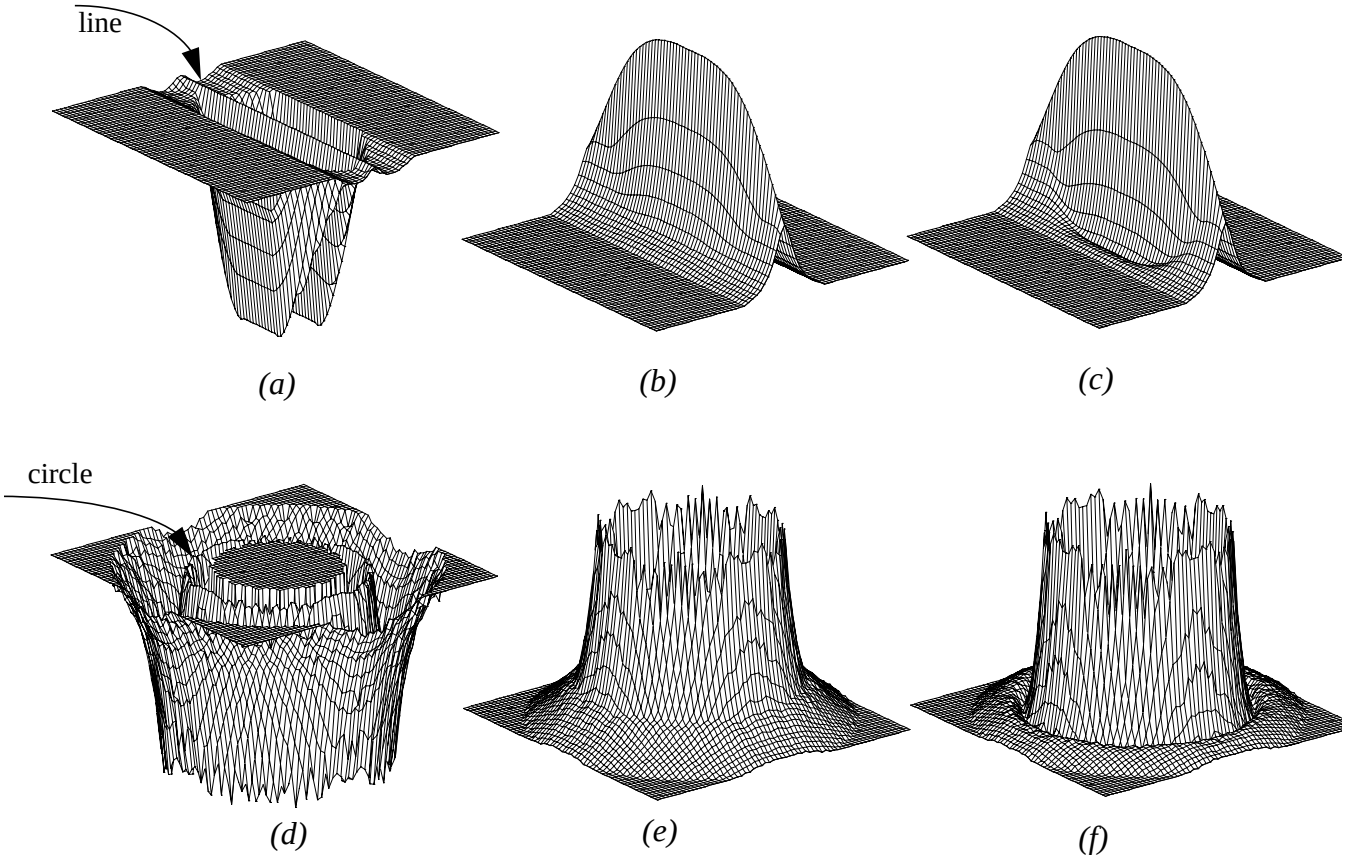


Figure 15 (a) Eccentricity only map of a saliency map of a straight line. (b) raw saliency map. (c) Product of (a) and (b). (d) -(f) same for a perfect circle.

is close to where the real line passes. In Figure 15(b) we see the raw saliency map of the same line. The Enhanced Saliency map is simply the product of the two maps, point by point, and will sharpen the edges of the correct curves, thus creating a non-maximum suppression affect(Figure 15(c)). Similar maps for a perfect circle are depicted in Figure 15(d-f).

The low eccentricity in the vicinity of the correct curve is due to the large variance of votes in these areas.

It may look at first sight that the field, while voting for the correct curve, contaminates the environment by voting for many other sites in the image. This can be regarded as noise, and is inherent to the process. However, while the fields voting *for* a curve agree along the curve, they disagree in any other area of the image. This means that the contribution of a complete curve to the environment is almost *iso-*

The product of our new eccentricity measure and the raw saliency measure λ_{max} yields the junction saliency operator:

$$E \cdot \lambda_{max} = (\lambda_{min} / \lambda_{max}) \cdot \lambda_{max} = \lambda_{min} \tag{4.9}$$

This process creates a *Junction Saliency map*. Interestingly enough, this map evaluates to just λmin at every site (as shown in (4.9)), which simply means that the *largest non-eccentric* sites are good candidates for junctions. By finding all *local maxima* of the junction map we localize junctions (see results in Figure 17).

Knowledge of the location of junctions is crucial in the linking phase of any edge detection scheme. As we show in the results, linking can be performed by following curves from one junction to the next.

4.4 Properties of the extension field

4.4.1 A longer line implies a stronger and more directed field, but up to a point

Using a simple example, we demonstrate the behavior of the field when extending a straight line. Figure 14 shows a cross-section of a saliency map computed on a series of straight lines with increasing lengths. Clearly, the saliency grows as a function of the length, and the map becomes more directed (thinner ridge). Also, saliency converges to some finite value which is just the infinite integral along the main axis (x -axis in Figure 6) of the Extension Field. This can be explained intuitively by noting that the influence of a given edge segment decreases monotonically with distance, so a given site can only gather votes up to a certain "influence radius". This observation can be used to estimate absolute saliency, since the maximum saliency value we expect (for a 'clean' input) is known. The maximum saliency will change according to the curvature of the curve, as the constraints dictate.

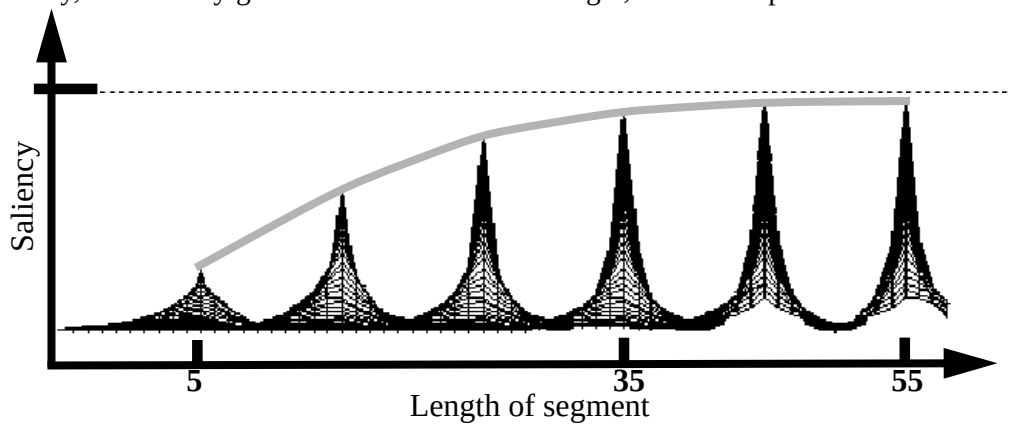


Figure 14 Saliency of a line does not grow forever. It converges to some value which is the infinite straight integral of the extension field.

Our intuitive notion of ‘agreement’, or of a majority vote on a continuous scale, is consistent with the above definition. This means that in all cases where we feel that collection *A* has better ‘agreement’ than collection *B*, the corresponding eccentricity values will share the same relationship (i.e. *E(A)>E(B)*). This is not to say that both functions are equal, but merely that both are monotonic.

Eccentricity values by themselves cannot perform as saliency measures since sites with very little voting strength can produce high eccentricity values. In fact, consider a site far away from where the ‘action’ is, which accepts exactly *one* vote (This can happen in practice). The eccentricity value is 1, but the site is of no importance.

However, Consider λ_{max} itself. Obviously,

$$\frac{\lambda_{min} + \lambda_{max}}{2} \leq \lambda_{max} \leq \lambda_{min} + \lambda_{max} \tag{4.6}$$

By (4.6) λ_{max} is bounded from both sides by the proximity measure in (4.5) *and* has the eccentricity coded into it. When the value leans towards the left side of (4.6), eccentricity is low and vice-versa.

Thus, **λmax is chosen as the raw saliency measure in our scheme.**

This choice however, may still amplify locations which are very strong in terms of number of votes, but weak in eccentricity[1]. The product of E and λ_{max} produces the desired result, termed the *enhanced saliency* measure SM, or:

$$SM = \lambda_{max} \cdot (1 - \lambda_{min}/\lambda_{max}) = \lambda_{max} - \lambda_{min} \tag{4.7}$$

Thus, λmax-λmin is chosen as the enhanced saliency measure.

It is important to note that other functions of the eigenvalues can also satisfy the same conditions of monotonicity, but the ones chosen seem to be the *simplest* possible indicators of the desired behavior.

Detection of Junctions

A junction is defined as a *salient* point having *a low* eccentricity value.

Regular (non-junction) points along a curve are expected to have high eccentricity values. On the other hand, junction points are expected to have low eccentricity, since votes were accumulated from several different directions. By combining the eccentricity and the eigenvalue at a point, we acquire a continuous measure of the likelihood of that site being a junction. We redefine our previous definition of eccentricity slightly, so that low eccentricity scores high, or:

$$(E = \lambda_{min}/\lambda_{max}) \Rightarrow 0 \leq E \leq 1 \tag{4.8}$$

1. For example, accumulation points and junctions! (where $\lambda_{min} \cong \lambda_{max}$)

This number, however, does not take into account the agreement between votes. A curve is salient if many votes (high λ_{max}) are collected, *and* their orientation is consistent (low λ_{min}). The eccentricity $1 - (\lambda min / \lambda max)$ is therefore a good measure of agreement.

When that value is multiplied by the previous raw saliency measure we achieve better selectivity, and only *curves* are highlighted. This results in a map defined by $\lambda max - \lambda min$.

Justification What we are looking for is a function that accepts positive vectors as input and results in a measure of the agreement in their orientation. The result should satisfy several criteria:

- We want the result to be normalized, so that we can compare different sites on a standard scale.
- The measure needs to be monotonically increasing with the addition of positive contributions.
- It should give higher values to 'better' (more directed) spatial arrangements of vectors.
- We want the effect of proximity to be independent to the effect of agreement.

It is easy to show how the model behaves when a single vector is added to it. Assume the variance-covariance matrix is as follows at state *t:*

$$C^t = \begin{bmatrix} m_{20}^t & m_{11}^t \\ m_{11}^t & m_{02}^t \end{bmatrix} \tag{4.2}$$

The sum of the eigenvalues is the trace of the matrix:

$$\lambda_{\min}^t + \lambda_{\max}^t = m_{20} + m_{02} \tag{4.3}$$

Now adding a new vector $V = [R\cos\theta, R\sin\theta]^T$ to the system will result in a new state *t+1*:

$$\lambda_{\min}^{t+1} + \lambda_{\max}^{t+1} = m_{20} + m_{02} + (R\cos\theta)^2 + (R\sin\theta)^2 = m_{20}^t + m_{02}^t + R^2 \tag{4.4}$$

Note that the angle θ has disappeared on the r.h.s. of (4.4). This means that the sum of eigenvalues is independent of the orientations of the voting vectors and can hence be used as an indicator of proximity (a wider sense of proximity of course), and as a naive saliency measure.

Equation (4.4) can be written as:

$$\lambda_{\min} + \lambda_{\max} = \sum_{i=1}^{N} R_i^2 \tag{4.5}$$

Where N is the number of segments in the original image.

We define the eccentricity $E = 1 - \lambda_{\min}/\lambda_{\max}$ as a measure of agreement. This value is between 0 and, Since $\lambda_{\min} \le \lambda_{\max}$, and are both non-negative.

The output of the directional convolution, *O(i,j)*, is in the form of a variance-covariance matrix for *each* point in space (see the next section for justification),

$$O(i,j) \;=\; m_{u,v}^{i,j} \;=\; \begin{bmatrix} m_{20}^{i,j} & m_{11}^{i,j} \\ m_{11}^{i,j} & m_{02}^{i,j} \end{bmatrix}$$

We are now ready to define the directional convolution operator $EF \oplus I \;=\; O$ of a vector field *EF* with an input field *I*,

$$m_{uv}^{x,y} \;=\; \sum_j \sum_i \left\| I_{i,j} \right\|^2 \cdot \left[\left(R_{\bar{I}_{i,j}} \cdot T_{i,j} \cdot EF \right)_x^{x,y} \right]^u \cdot \left[\left(R_{\bar{I}_{i,j}} \cdot T_{i,j} \cdot EF \right)_y^{x,y} \right]^v$$

where $0 \le u, v \le 1$ and $u + v \;=\; 1$. This will define all elements in *O(x,y)*.

When the input is a scalar field, the rotation operator becomes the identity operator and has no effect.

In order to obtain a majority vote regarding the preferred orientation of a given position, we treat the contributions to a site as being vector weights, and compute moments of the resulting system. Such a physical model behaves in the desired way, giving both the preferred direction and some measure of the agreement. We use the direction of the principal axis ($EVmin$) of that physical model as the chosen orientation (See equation (4.1)).

$$\begin{bmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{bmatrix} \;=\; \begin{bmatrix} EVmin \\ EVmax \end{bmatrix} \begin{bmatrix} \lambda min & 0 \\ 0 & \lambda max \end{bmatrix} \begin{bmatrix} EVmin^T & EVmax^T \end{bmatrix} \tag{4.1}$$

This acts as an approximation to the desired majority vote, without the need to consider the individual votes, but rather the statistics of the set.

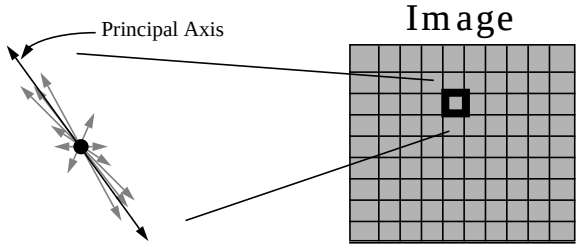


Figure 13 The principal axis of the votes collected at a site is taken as an approximation of the preferred direction.

4.3 Vote interpretation

At each pixel location, we now have the result of a vector vote in the form of an ellipse (λ_{min}, λ_{max}, θ). This information needs to be interpreted.

Clearly, λ_{max} is the raw saliency measure as it relates to the number of votes collected.

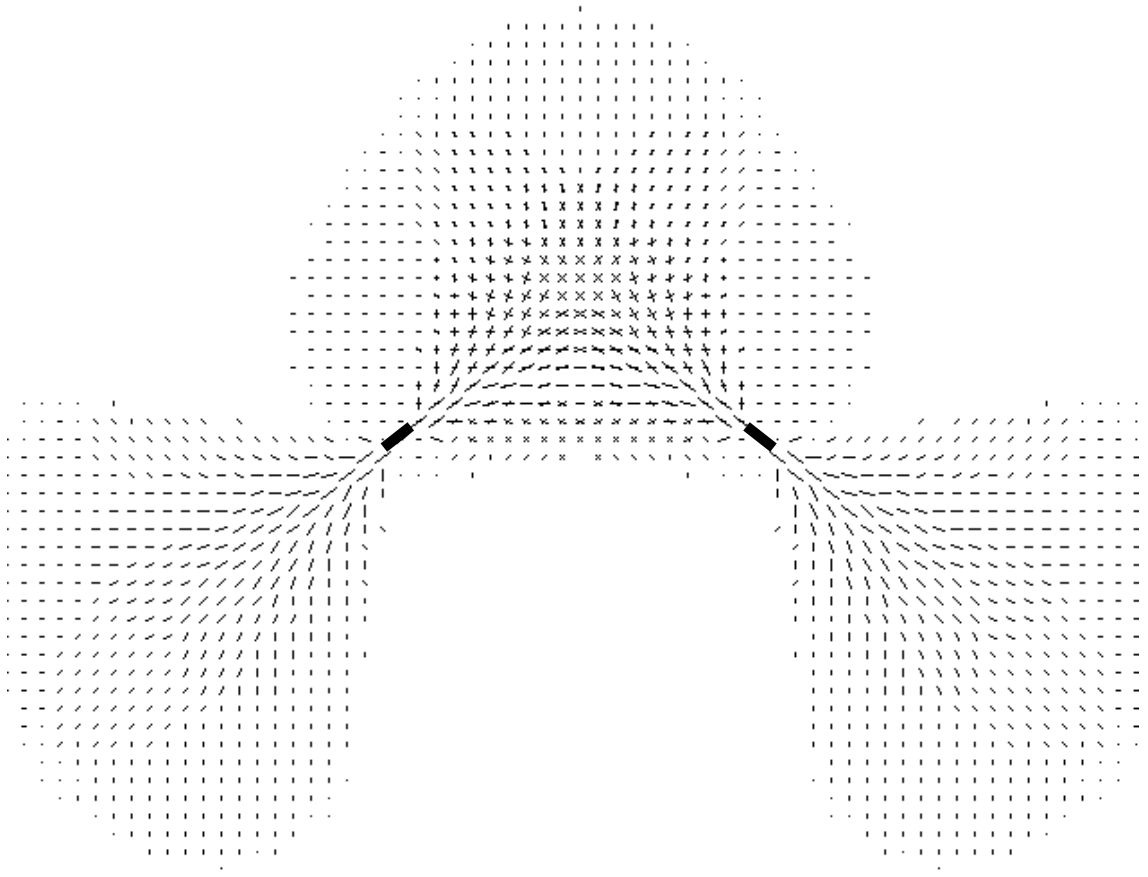


Figure 12 Superposition of two Extension Fields over a scenario of two co-circular short segments (dark lines).

For *R*, we first define a rotation matrix ($M_{a,b}$), that brings the unit vector $\{0,1\}^T$ to any given unit vector $\{a,b\}^T$. Such a matrix always exists, but may not be unique. This presents no problem, since all our fields are symmetric around the x-axis, and as such are invariant to the specific rotation. Thus, for *R* we get,

$$R_{a,b} \cdot \bar{V}(i,j) \;=\; M_{a,b} \cdot \bar{V}(M_{a,b} \cdot \{i,j\}^T)$$

Note that for the rotation operator, first the coordinate system has to be transformed, and *then* the actual vector at that point in space needs to be rotated.

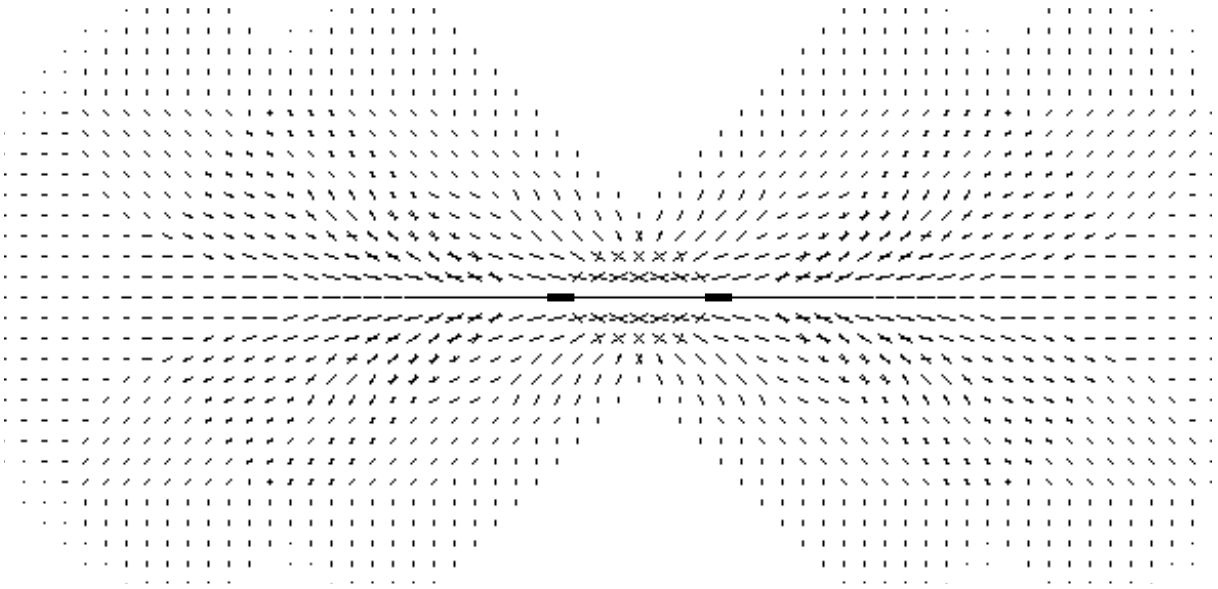


Figure 11 Superposition of two Extension Fields over a scenario of two co-linear short segments (dark lines).

The next step is to represent these accumulated vector votes at each pixel in a convenient and meaningful way. We argue that the relevant information is computed simply by the distribution of these votes, and specifically well approximated by the best fit ellipse representing the moments of these votes. We give the mathematical details below.

4.2 Vote representation

The strength of the field is scaled by a measure of the contrast of the segment, so that stronger segments have stronger votes throughout the space.

Note that, although the process is local in essence, the fields impose some global order, and one line segment can implicitly 'vote' for a large curve without any *explicit* global reasoning involved.

As before, we define the Extension Field, *EF(i,j)* at any site*(i,j)*, to be a vector field:

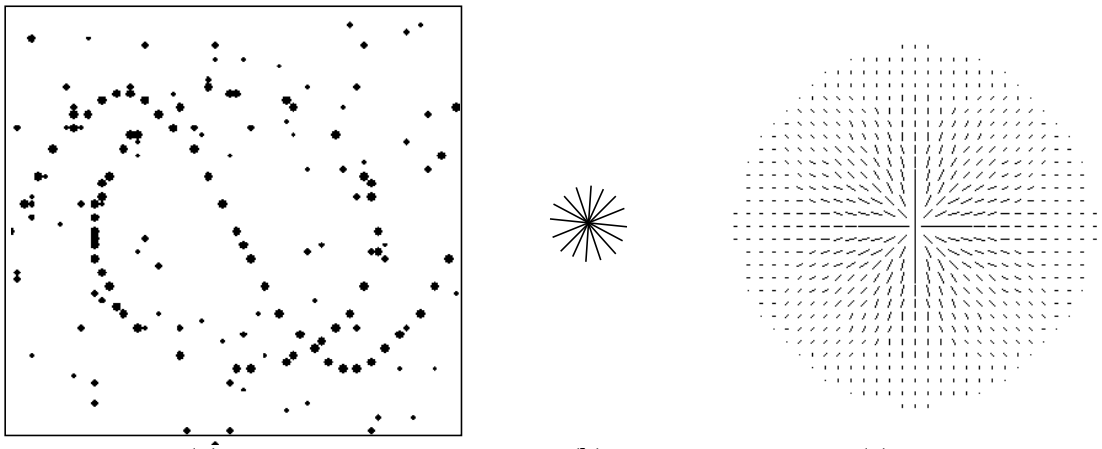
$$\overline{EF}(i,j) \;=\; (EF_x, EF_y)^T$$

The input *I*, will be denoted with a bar when oriented (a vector field), and without when non-oriented (a scalar field).

We define the translation (*T*) and rotation (*R*) operators to translate and rotate an arbitrary vector field (*V*) respectively. For *T* we have,

$$T_{\Delta x,\,\Delta y} \cdot \bar{V}(i,j) \;=\; \bar{V}(i+\Delta x, j+\Delta y)$$

which simply translates the indices of the target vector field.



(a) *(b)* *(c)*

Figure 9 (a) A typical dot formation. (b) A multi-directional edge and the resulting point field.(c)

is constructed from a degenerated instance of an ellipse (a line), while the Point field is created from a 'circular' edge point. This makes full use of our input model and allows treatment of mixed images[1] in a consistent and unified way (see Figure 10).

In practice, in was found that the Extension field, although designed for perfect orientations, tolerates some degree of uncertainty in the input samples. This is due to the special interpretation of the votes, which is described later.

Figure 10 A continuum of field constructors. From the maximum certainty in orientation (left), to maximum uncertainty (right).

4 Computation of the Saliency Map

4.1 Vote Accumulation

Having designed the field for a given input site (from a dot to an edgel) we superpose these votes from each active site at each location. A pixel receives a vote from *all* sites whose field overlap this pixel, and each vote has the form of a vector, given by its strength (length) and orientation.

For each voting site, we align the appropriate extension field with the direction of the site, center the field at the site, and compute the vector contribution at each location covered by the field. The is repeated for all input sites. This process is similar to a convolution with a mask, except that the output is a vector instead of scalar. The procedure is illustrated in Figures 11 and 12.

1. i.e. having varying certainty measures (or a mixture of dots and lines)

eral correspond to any analytical function. For practical reasons, however, we have decided to 'guess' an approximating analytical function to the Extension Field (EF) (equation 3.2).

$$\overline{EF}(x,y) = \begin{cases} e^{-Ax^2}(1,0)^T & , \quad y=0 \\ \left(e^{-(Ax^2 + B\,\mathrm{atan}(|y|,|x|)^2)}\right)\left(\frac{x}{R}, \frac{y}{R} - 1\right)^T & , \quad y \neq 0 \end{cases} \qquad (3.2)$$

$$\text{where} \quad R(x,y) = \frac{\left(x^2 + y^2\right)}{2y}$$

The parameter A controls the proximity decay, while B controls the decay due to higher curvature. The parameters A and B were selected empirically based on the above-mentioned constraint. They are not however, independent, and changing one will require to change the other. The shape is thus that of a decaying exponential, where the decay is a function of both the distance and the radius of the corresponding circular arc. The parameter A is related to the size of gaps we are able to recover, and should be derived from image size and sparseness of data. In our implementation, A was selected to bridge gaps of up to 50 pixels (A=0.003), and B was adjusted to loosely satisfy the above-mentioned constraint (B=2.85).

Note that the values of A and B need to be tuned only once during the design of the field, and are kept constant through all our experiments.

3.2 Other fields

3.2.1 The Point Field

A dot image is a degenerate case of an edge image, where the edges have no direction. Such maximum orientation uncertainty in the input fits our input model well, and allows us to handle such cases in a uniform way. Obviously, perception is weakened by the loss of orientation data, and we are only able to handle cases with a moderate amount of noise. The only applicable perceptual law is that of *proximity.*

A suitable field must have circular symmetry, and in practice is constructed by convolving our original extension field with a 'multi-directional' edge segment (Figure 9). A typical input is shown in Figure 9, where a broken sine wave and a set of random points on a circle are embedded in noise.

Another scenario where the point field could be useful is in images where co-curvilinear formations between features other than dots are present. In such cases we would like to treat the image as if it was made out of non-directional tokens (or dots), and apply the point field to it.

3.2.2 Unifying the field concept

We unify the maximum certainty (Extension) field and the Point field (and all fields in between) by considering a continuum of eccentricities associated with the multi-directional edge. The Extension field

low the main diagonals. This merely means that we choose *not* to vote for pixels in that area, and additional 'in-between' information is necessary to reconstruct curves between such pairs.

Another reason all values beyond the two main diagonals are zeros, is a technical one. Having a segment vote for a point in space which is more than 90 degrees away (along a circular arc) could potentially cause unrelated segments to vote for the same curve, even though such a curve should not connect them.

Strength The main consideration here is that we would like a decay due to distance and a a decay due the higher curvature. Many functions can fulfill this, and the problem is under-constrained. We add the following constraint. The best way to determine these values is by considering an *intentionally* ambiguous or *undecidable case*. The assignment of actual probabilities to the field is thus performed as follows: We consider two short edge segments, *perpendicular* to each other and apart[1] (see center of Figure 8). This scenario, we claim, is a middle point between a clear choice of a connection by a sharp

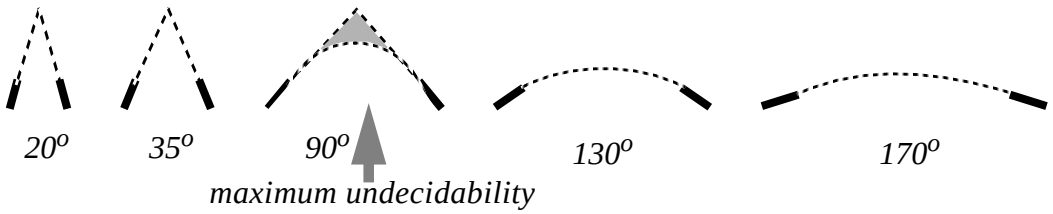


Figure 8 5 arrangements of two segments, each with a different separation angle. Angles much smaller than 90 degrees suggest a corner, while angles much larger than 90 degrees suggest a smooth connection.

junction and a connection by a smooth curve. We regard this to be a most competitive scenario in terms of grouping of the two line segments. We thus assign probabilities to the field elements in such a way that all paths connecting these segments are assigned roughly the *same* saliency, and there *does not* exist any single best path between the two. More precisely, we set the field element strengths such that all values within the marked triangular region are the same. Such a scenario, when repeated for all distances, removes all but one degree of freedom as to the choice of values for the field. The weights will not in gen-

1. This scenario is termed '*the maximum undecidability arrangement*'.

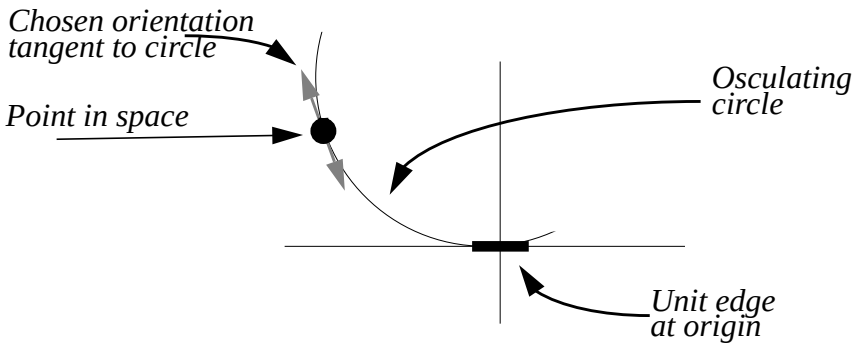


Figure 7 Assigning a direction for every point in space

tween compatible edgels before. The circular vote here is somewhat different, since the field is *not* designed to evaluate pair-wise compatibility, but rather to 'suggest' plausible continuations everywhere in space.

In trying to computationally evaluate the various constraints over a given curve we find that a (somewhat revised[1]) measure of Total Curvature (as used by [22]) encompasses most of the desired constraints. We define the Total Curvature (TC) to be:

$$C = \int_{\Upsilon} \left| \frac{d\theta}{ds} \right|^{\alpha} ds, \qquad \alpha > 1 \tag{3.1}$$

This is the integral of the curvature (brought to some power) along the curve, where θ is the tangent along the curve γ. The variable α is traditionally taken to be equal to 2, but it can be shown that the choice of a circle as the connecting curve in the scenario shown in Figure 7 minimizes the TC for all values of α greater than 2 [8]. A larger α would just penalize sharp turns more than a smaller one. We thus set the orientations of the field elements to be tangent to a circular arc connecting the origin and that point in space.

It is important to note that two oriented points *cannot* always be joined by a circular arc. Such pairs may be subject to some more complicated smooth curve connection, but we have decided not to handle explicitly these cases. It is quite obvious that extending a curve beyond the 90 degree point does *not* satisfy the minimum curvature constraint.[2] For this reason our Extension field has zero values above and be-

1. In [22], α=2 is used, and the absolute value is redundant.
2. Note that our proof holds only in the 0 to 90^0 interval. It is sufficient to give the elliptic continuation as a counter-example to show that the behavior beyond the 90 degree point is different.

promise between extension fields is also central to the approach, and will be presented in a more formal way later.

It is not reasonable to expect that extension curves from two different extension fields will align throughout their extent. It is more likely that such extensions align locally in many places. For that reason, the extension field will consist of local best candidates for extensions. In the next section, we define the exact shape and usage of the Extension Field.

3 The Extension Field: design and implementation

Definition: An ***Extension Field*** is a maximum likelihood directional vector field describing the contribution of a single unit-length edge element to its neighborhood in terms of length and direction.

In other words, it votes on the preferred direction and the likelihood of existence of every point in space to share a curve with the original segment. The field is of *infinite* extent, although, in practice, it disappears at a predefined distance from the edge. Figure 6 depicts the Extension Field.

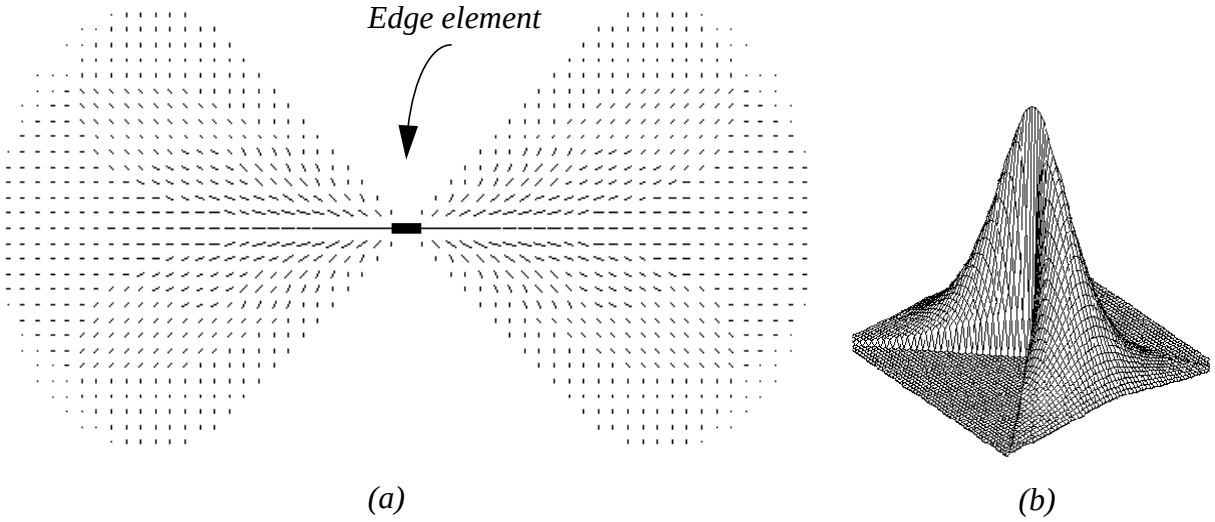


Figure 6 The basic Extension Field. (a) Direction, and (b) Strength.

The design of this field needs to account for the shape of the field, the orientation at each site, and the strength at each site. The first two aspects are quite straightforward, while the last needs more explanation.

3.1 Design of the Extension Field (Orientation and Likelihood)

Shape and Orientation. Since we favor small and constant curvature, the field direction at a given point in space is chosen to be tangent to the osculating circle passing through the edge segment and that point, while its strength is proportional to the radius of that circle. Also, the strength decays with the distance from the origin (the edge segment). The choice of a circular extension agrees with the constraint of smallest total curvature. We note that Ullman [25] proposed the usage of the a circular connection be-

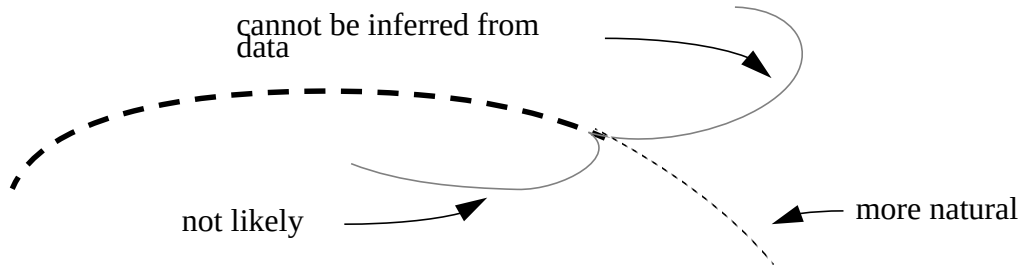


Figure 4 What is the shape of the most 'natural' extension to a given curve?

We would thus want to consider *all* the curves in such a way that the extension is smooth, influenced most by the behavior of close-by points of the curve, but can assume any form. Note that such extension can be constructed even if the curve is fragmented, and is robust, being the 'average' of many segments.

We go further than that: Rather than having only the best extension, we would like to list *all* possible extensions in the order of their likelihood. This suggests the use of some kind of a *field* (or flow) emanating from the end-points of a segment. The idea of a *field* plays a major role in our scheme, as we will show later.

2.1.3 Best Connection between Two Line Segments

Another instance of over-constraining arises when two *subparts* of two curves form a natural connection, which would require a fit-then-connect approach to solve a difficult segmentation problem.

The situation occurs whenever a 'compatibility figure'[1] of two close-by segments is computed in order to decide whether these segments should be grouped together. When determining the compatibility of two lines we would like to consider for each line its best extension (or extension field, as discussed in 2.1.2) and arrive at some compromise as to the best path between them. In other words, each of the curves

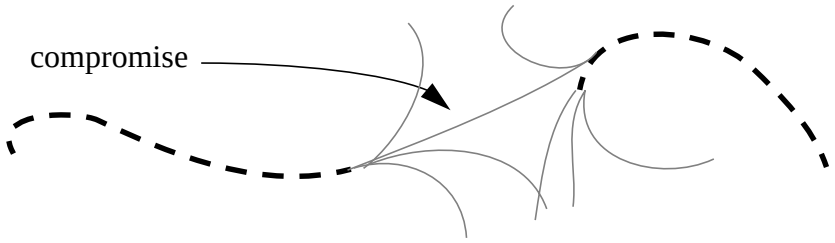


Figure 5 What is the best path between these two curves?

votes (using its field) for a family of curves. If the curves should really be connected then some extension from curve 1 would *align* with another extension of curve 2, to form the *compromise*[2]. The idea of a *com-*

1. A measure of 'agreement' between two curves, based on the difference of the end-point tangents and separation. (as used by [5, 14, 15])
2. Which is not necessarily the *best* continuation of any of the existing curves, or even a connection between the *given* end-points.

2.1 Rationale for the Extension Field

In order to define saliency qualitatively, we start by writing down the major constraints which govern our mechanisms of saliency.

2.1.1 The perceptual constraints

Our underlying goal is to keep the interpretation as simple as possible (in the 'Gestalt' sense). This translates into four major constraints:

1) Co-curvilinearity - In the lack of other cues, smooth continuation (co-curvilinearity) is the only interpretation.
2) Constancy of curvature - We tend to extend a curve of some constant curvature with the same curvature, keeping the interpretation as simple and regular as possible, yet consistent with our sensory information. This principle is called Prägnanz by Gestaltists (see Figure 3).

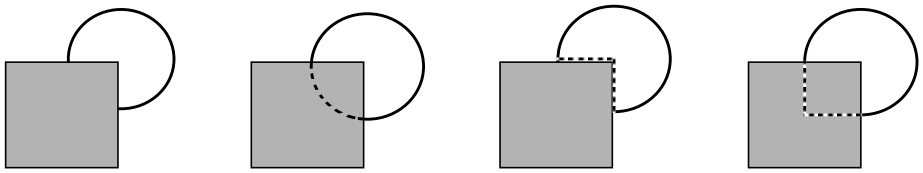


Figure 3 An obscured figure (a) triggers the perception of simple shapes (b), instead of the more complex (c) and (d). (From [20])

3) Favoring low curvatures over large ones - Humans seem to connect fragmented line segments in a way that the increase in total curvature is minimum (see Ullman and Sha'ashua[22]).
4) Proximity - Closer segments influence each other more than distant ones.

The above mentioned constraints operate on **curves** (with the possible exception of 4), whereas our input consists of first order approximations of these curves. We therefore need to both infer curves, *and* connect them. In our scheme, this circular problem is overcome by solving these two problems concurrently. This way, we avoid making hard choices as to the initial curves, which would then over-constrain the possible extensions.

2.1.2 Extending a Curve

One example of such a strong constraint is in the use of the tangent to the end-point of a curve. Given a line segment we ask the question: What is the shape of the most 'natural' extension, based on the mentioned constraints?

Many approaches in the past (e.g. [5,14]) used the *tangent* of the end-point to determine the best extension. This approach cannot always work properly for three reasons:

1) The tangent is very sensitive to noise and may introduce large errors.
2) The end-point may not be determined uniquely if the curve in question is fragmented.
3) The extension can only be a straight line (first order), thus not taking into consideration the global shape of the curve.

In our method, *every* site (pixel or other cell) collects votes from *each* segment in the image. These votes contain orientation and strength information preferred by the voting segment. A measure of 'agreement among votes' (in terms of orientation) is computed, and sites which have high agreement values are considered salient. In more technical terms, a vector field[1] is generated by each segment, and a function over the whole space determines points of saliency. A subsequent step links areas of high saliency to produce a description in terms of curves and junctions.

Our voting scheme is somewhat related to the Hough transform approach [11], but can detect shapes defined by their *properties* (smoothness etc.) rather than by their *exact shape* (lines, circles, etc.). A study of the relations between our scheme and the Hough Transform is given in the section "Comparison With the Hough Transform" on page 25

The process is likely to produce features more similar to what we expect, both in terms of saliency and connectivity. Also, since noise is not likely to produce high 'agreement' values, it becomes attenuated.

We believe that the use of a non-linear global voting scheme is fundamentally different from the many local techniques reported before. Such a non-linear *global* approach is difficult to decompose into an iterative *local* scheme.

Our system can take as input a set of oriented features, as produced by a typical edge-detector, as well as features with an associated directional uncertainty, and even edges with *no* associated orientation (dot formations). The output of our system is a set of oriented features, with an associated strength reflecting its saliency, and an uncertainty of orientation.

It is convenient to represent each site (input or output) by an ellipse, whose major axis points in the direction of the edgel direction, whose length is proportional to the strength, and whose minor axis reflects the associated uncertainty. This is illustrated in figure 2. An isotropic feature (such as a dot) would thus be represented by a circle.

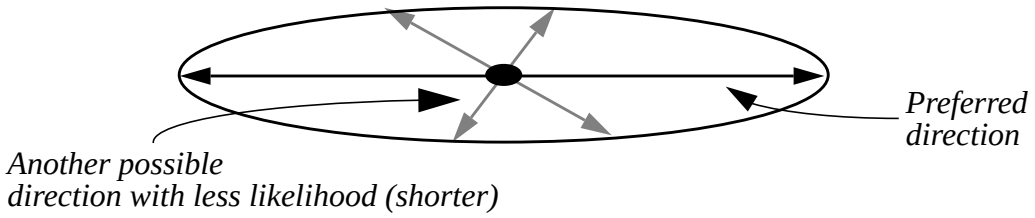


Figure 2 Input and output site model. Every site is associated with a preferred direction, strength and eccentricity (or uncertainty).

1. Which we later call the **Extension Field.**

	Lowe[14]	Ahuja & Tuceryan [1]	Dolan & Weiss[5]	Mohan & Nevatia [16]	Ullman & Sha'ashua [22]	Parent-Zucker [17]	Heitger & von der Heydt [9]	Our scheme
Operator	Local	Local	Local	Local	Extensible (local)	local	Global	Global
Primitives	Straight lines	Dots	Straights	curves	straight lines and curves	curves	endpoints and T-junctions	dots, lines and curves
Control	One pass	Iterative	Iterative	Relax-ation Pro-gressive	Parallel-pro-gressive (it-erative)	relaxation	one pass	One-pass convolution
Noise im-munity	Not clear	Good	Good	Good	Moderate	Good	Not han-dled	Very good
Scale	One	One	Hierarchy	One	One	One	One	One
Parameters	Yes	No	Yes	Yes	Yes	Yes	Yes	None
Pre-atten-tive (Do-main free)	Yes	Yes	Yes	No (yes)	Yes	Yes	Yes	Yes
Special fea-ture	First	Dot cluster-ing. parame-ter free	Multi-res-olution	Symme-try, high-level con.	Saliency map	local ker-nels		Saliency map, unified, parameter-free
Sensitive computa-tions	Yes	No	Yes?	Yes	Yes	Yes	No	None

Table 1: Comparison of different grouping techniques

2 Overview of Our Approach

As was presented before, the physical evidence extracted *locally* from images (e.g. through edge detectors) is in many cases ambiguous and does not correspond to the expected perception of the image. It is thus necessary, we believe, to impose *global* perceptual considerations even during low-level processing. These constraints need to be generic, as they should embody expected properties of all scenes.

We briefly explain below what is computed, how it is computed, and give a rationale for the approach.

In signal edge-detection, we extract edges and assign position, orientation, and strength based on local physical measurements. Unfortunately, the relationship between the strength of the perceived edges and the strength of the measured signal is not straightforward. We therefore propose to generate a different measure of strength, which we call saliency, based on more global expected properties of the observed scene, such as co-curvilinearity. This idea has been tried in the past, but is rather difficult to implement efficiently, and in a robust way.

Lowe [14] discusses the Gestalt notions of *collinearity, co-curvilinearity* and *simplicity* as important in perceptual grouping. Ahuja and Tuceryan [1] suggest methods for clustering and grouping sets of points having an underlying perceptual pattern.

Dolan and Weiss [5] demonstrate a hierarchical approach to grouping relying on compatibility measures such as proximity and good continuation. Mohan and Nevatia [15] assume a-priori knowledge of the contents of the scene (i.e. aerial images). A model of the desired features is then defined, and groupings are performed according to that model. In a later work [16], groupings based explicitly on symmetries are suggested, but the first connectivity steps are performed locally.

Ullman [25] has suggested that a curve joining two edge fragments is formed by the pair of circular arcs which minimize the integral of curvature squared. He also proposed a network model, but no results were shown. Clearly, elliptical curves cannot be constructed by joining only a pair of circular arcs. Also, this scheme cannot be easily generalized to a set of 3 or more edge fragments, and does not allow for outliers. In a sense, what we propose next is an extension of the same idea where a curve is formed (and/or approximated) by joining an unlimited number of (possibly) short circular arcs.

Parent and Zucker [17] describe a relaxation labeling where local kernels are used to estimate tangent and curvature. These kernels use support functions based on co-circularity. Somewhat similar kernels are used in our scheme, but applied in a different way.

Heitger and von der Heydt [9] make use of anisotropic selective filters which are combined pairwise to recover occluding contours. The scheme takes as input endpoints and T-junctions, and results in the most natural connections of those. It is based on neurophysiological observations, but does not handle noise, and assumes that endpoints and T-junctions are available by some other means. The authors present convincing results on real images.

Sha'ashua and Ullman [22] suggest the use of a saliency measure to guide the grouping process, and to eliminate erroneous features in the image. The scheme prefers long curves with low total curvature, and does so by using an incremental optimization scheme (similar to dynamic programming).

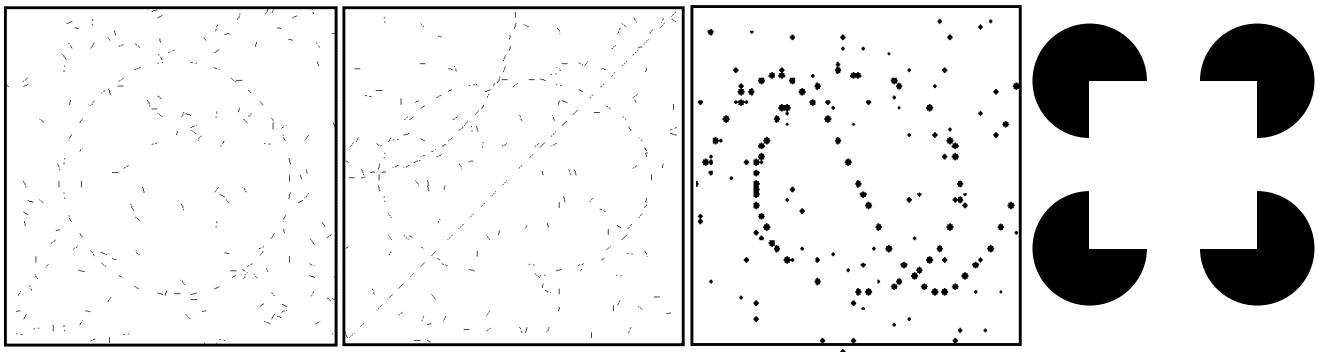
Others ([30,18]) have also looked at similar problems.

The main features of some of the more important works are summarized in table 1, and contrasted with our scheme. It is interesting to note that virtually all proposed algorithms use *local* operators to infer more global structures. Also note that many of the schemes are iterative, relying on one relaxation (or minimization) scheme or another, and are similar in that sense. The main differences are in the choice of the compatibility measures or the function to minimize.

tension Field. We follow by a comparison with the classical Hough transform and conclude by offering a set of other fields for special purposes.

1.1 Perceptual Grouping

Perceptual Grouping refers to a class of visual phenomena where clustering of physically non-connected elements in the image occurs. This task is equivalent to a figure-ground discrimination when patterns are embedded in noise. Figure 1 depicts examples of perceptual groupings which are of interest to



(a) *(b)* *(c)* *(d)*

Figure 1 (a) & (b) Two instances of perceptual arrangements. (c) A dot formation. (d) The Kanizsa Square.

us, and are considered to be the result of a pre-attentive process. Such processes are known to take several hundreds of milliseconds (200-500 ms) to complete, and are thus not likely to utilize any high-level reasoning mechanism [3].

The circle in the middle of figure 1(a) is easily distinguishable from its noisy background. Furthermore, we tend to fill the gaps, or more precisely, we are able to complete the circle mentally. The same holds for the geometrical patterns in 1(b). Figure 1(c) depicts a dot formation. Again, grouping of certain dots is possible, and salient curves are noticeable. A more striking example of *illusory* contours is found in the Kanizsa illusion [12] shown in Figure 1(d). Here we perceive edges which have no physical support whatsoever in the original signal.

The Gestalt psychologists (e.g. [28,3]) were among the first to address the issues of pre-attentive perception. Many ‘laws of grouping’ were formulated, but none put in any algorithmic language. Furthermore, the rules tend to supply conflicting explanations to many stimuli. This makes the computational implementation of such laws nontrivial. With input in the form of edges, the laws most relevant to our work relate to *proximity* and *good continuation*.

The same constraints can be derived using different arguments. For instance, Lowe [14] uses a probabilistic framework to derive similar constraints. We now briefly review some of the related literature, and provide a comparison table.

Inferring Global Perceptual Contours from Local Features

Gideon Guy and Gérard Medioni

Institute for Robotics and Intelligent Systems
University of Southern California
Los Angeles, California 90089-0273

1 Introduction

Computer vision can greatly benefit from perceptual grouping. Perceptual Grouping can be classified as a mid-level process directed toward closing the gap between what is produced by state-of-the-art low-level algorithms (such as edge detectors) and what is desired as input to high level algorithms (perfect contours, no noise, no fragmentation, etc.). Many researchers resort to using synthetic data as their input because of this weak link.

Determining groups in a given set of points or edgels can be a very expensive computational task. Not only does one need to consider all possible sub-sets for grouping (a complexity of $O(2^{\#\_of\_points})$), but the actual measurement of compatibility of a sub-set is not well defined. Methods for edge labeling (such as [4,27]) assume perfect segmentation and connectivity, and define constraints which are only valid under these assumptions. These methods cannot work on 'real' edges. Other methods, like shape from contour [26], and [21, 24], also rely heavily on the connectedness of the edges, and can benefit from the removal of noise (erroneous segments). Pattern recognition schemes (such as[23]) rely on (at least) partial connectedness of the edges, and cannot function if the edge image is very fragmented. Also, the amount of noise is directly proportional to the computational cost of finding 'real' objects in a scene. In addition, the complexity of most pattern recognition schemes is directly proportional to the number of distinct primitives in the input.

Using global perceptual considerations when attempting to connect fragmented edge images can alleviate many of the above problems, as we show here.

We introduce a general algorithm capable of highlighting features due to co-curvilinearity and proximity. We suggest the **Extension Field** as a 'voting pattern' representing a large family of smooth curves, all at once. We explore the properties of the Extension Field, and suggest specialized fields that can be used within the same computational paradigm to reveal perceptual phenomena such as endpoint formations and straight lines. Experimental results with different types of distortion applied to the input are also presented. We start by explaining the original algorithm, then describe some of the properties of the Ex-